MULTIPERSPECTIVE RECOGNITION APPLIED TO THE COMPUTER-AIDED MEDICAL DIAGNOSIS – A COMPARATIVE STUDY OF METHODS

Marek W. Kurzynski, Edward Puchala

Wroclaw University of Technology, Faculty of Electronics, Wyb. Wyspianskiego 27, 50-370 Wroclaw, Poland

Abstract – **This paper deals with the multiperspective recognition technique applied to the computer-aided decisions in medicine. For three different concepts of multiperspective classification, i.e. direct, decomposed independent and decomposed dependent approach, several decision algorithms are presented. They are: probabilistic (empirical Bayes) algorithm, nearest neighbour algorithm, fuzzy method and artificial neural network of the back propagation and counter propagation types. Proposed methods and algorithms have been applied to the computer-aided diagnosis of chronic renal failure and decisions in non-Hodgkin lymphoma. Results of experimental investigations on the real data and outcomes of the comparative analysis of discussed algorithms are presented.**

Keywords – **Multiperspective recognition, probabilistic methods, neural network, fuzzy approach, chronic renal failure, non-Hodgkin lymphoma.**

I. INTRODUCTION

One of the most often used instruments for the computer-aided medical diagnosis are pattern recognition methods. This is because the problem of classification of a given pattern is identical to the task of medical diagnosis and therefore the methods and algorithms found in this theory are directly employed.

The classical pattern recognition problem is concerned with the assignment of a given pattern to one and only one class from a given set of classes. Multiperspective classification (MC) problem, which idea has been presented in [1] refers to a situation in which an object undergoes several classification tasks. Each task denotes recognition from a different point of view and with respect to different set of classes. For example, such a situation is typical for compound medical decision problems where the first classification denotes the answer to the question about the kind of disease, the next task states recognition of the stadium of disease, the third one determines the kind of therapy, etc.

This work presents several different approaches to algorithmization of the multiperspective diagnosis task and the implied decision algorithms. The following approaches will be explored in turn: the sample-based probabilistic approach, the nearest neighbour algorithm, the fuzzy system and artificial neural networks. It should be noted that for each of the four cases a set of diagnostic algorithms will be proposed. The way and scope of effective usage of the available data will be different for each set, as will the formal model for the corresponding decision task be. We will also perform an experimental comparative analysis of the proposed algorithms for real data that concern the diagnosis of chronic renal failure and decision problems in non-Hodgkin lymphoma. The performed experimental research will enable us to evaluate the presented conceptions along with usefulness of the proposed algorithms, as well as recommend them for possible practical applications.

II. MULTIPERSPECTIVE RECOGNITION

Let us consider N-perspective recognition problem. Let $x_k \in \mathsf{X}_k$ and $j_k \in \mathsf{M}_k$ denote the vector of features and the class number for the k-th recognition task, respectively.

In the case of MC we can in different manner define the action of recognizer which leads to the following recognition concepts [2]:

Direct approach: a decision rule Y of MC maps the space of entire feature vector X to the product $\mathsf{M}_1 \times \mathsf{M}_2 \times ... \times \mathsf{M}_N$. It means that now MC is a single activity in which a given pattern is classified into classes simultaneously for all recognition tasks.

Decomposed independent approach: an object is classified in a succesive manner for particular recognition tasks and the decision algorithm Y_k for k-th classification uses features x_k only, viz: Y_k: $\mathsf{X}_k \to \mathsf{M}_k$. This concept denotes decomposition of MC problem on N independent decisions which do not take into account the fact that they state steps of the whole compound decision process.

Decomposed dependent approach: an object is classified in a successive manner for particular recognition tasks and now the decision algorithm Y_k for the k-th classification uses not only features x_k but additionally the results of previous classifications, viz., $Y_k : \mathsf{X}_k \times \mathsf{M}_1 \times \mathsf{M}_2 \times ... \times \mathsf{M}_N \to \mathsf{M}_k$, This concept similarly denotes the decomposition of the whole MC problem but now the successive classification tasks are mutually dependent and fully describe real situations occurring in the computer-aided medical diagnosis tasks.

In next section we present different approaches to the construction of decision algorithms of multiperspective recognition (diagnosis).

III. METHODS

A. The Probabilistic Algorithms

Now we assume, that the vector of features x_k and the class number j_k (k=1,2,...,N) for the k-th recognition task of the pattern being recognized are observed values of random variables $\boldsymbol{x_k}$ and $\boldsymbol{j_k}$, respectively. Probability distribution of $\boldsymbol{x_k}$, $\boldsymbol{j_k}$, k=1,2,...,N is given by *a priori* probabilities p_{jk} and

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25 Oct 2001	N/A	-

Title and Subtitle Multiperspective Recognition Applied to the Computer-Aided Medical Diagnosis - A Comparative Study of Methods	**Contract Number**
	Grant Number
	Program Element Number
Author(s)	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Wroclaw University of Technology Faculty of Electronics Wyb. Wyspianskiego 27, 50-370 Wroclaw, Poland	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom.	
Abstract	
Subject Terms	
Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UU
Number of Pages 4	

class-condintional density functions (CDF) $f_{jk}(x_k)$. Using the Bayes decision scheme [3], we can receive the following recognition algorithms for considered concepts of multiperspective diagnosis:

Bayes algorithm for direct approach (Bayes-A)

$$y^{(BA)}(x) = (i_1, i_2, ..., i_N) \quad (1)$$

$$\text{if } \sum_{k=1}^{N} p(i_k / x) = \max_{i_1^{'}, i_2^{'}, ..., i_N^{'}} \sum_{k=1}^{N} p(i_k^{'} / x),$$

where *a posteriori* probabilities $p(i_k / x)$ can be calculated from the Bayes formula.

Bayes algorithm for decomposed independent approach (Bayes-I) (for *k*-th task)

$$y^{(BI)}(x_k) = i_k \quad \text{if} \quad p(i_k / x) = \max_{i_k^{'}} p(i_k^{'} / x) \quad (2)$$

Bayes algorithm for decomposed dependent approach (Bayes-D) (for *k*-th task)

$$y^{(BD)}(x_k, i_1, i_2, ..., i_{k-1}) = i_k \quad \text{if} \quad (3)$$

$$p(i_k / x_k, i_1, i_2, ..., i_{k-1}) = \max_{i_k^{'}} p(i_k^{'} / x_k, i_1, i_2, ..., i_{k-1}).$$

In the real world there is often a lack of exact knowledge of *a priori* probabilities and CDFs. For instance, there are situations in which only a learning sequence

$$S_L = (x^1, j^1), (x^2, j^2), ..., (x^L, j^L), \quad (4)$$

$$x^k = (x_1^k, x_2^k, ... x_N^k), \quad j^k = (j_1^k, j_2^k, ..., j_N^k),$$

that is a set of case records with the firm diagnosis is known. Now conceptually simple method is to estimate unknown *a priori* probabilities and CDFs and then to use these estimators to calculate *a posteriori* probabilities.

B. Nearest Neighbour Algorithms

Nearest Neighbour (NN) algorithm is very well known classification technique [3] using learning set (4). Its modified versions for three different concpts of MC, can be formulated as follows.

Direct multiperspective NN algorithm (NN-A)

1. Measure features x of the pattern to be recognized,
2. Find the nearest neighbour (in the space $\mathcal{X}$) to x from among learning patterns (let say x^m),
3. Assign the recognized pattern to the sequence of classes j^m.

Decomposed independent multiperspective NN algorithm (NN-I) (for the *k*-th task)

1. Measure features x_k of the pattern to be recognized,
2. Find the nearest neighbour (in the space $\mathcal{X}_k$) to x_k from among learning patterns (let say x_k^m),
3. Assign the recognized pattern into the class j_k^m.

Decomposed dependent multiperspective NN algorithm (NN-D) (for the *k*-th task)

Notation: $i_1, i_2, ..., i_{k-1}$ - results of recognition for the first k-1 tasks.

1. Measure features x_k of the pattern to be recognized,
2. Find the nearest neighbour (in the space $\mathcal{X}_k$) to x_k from among learning patterns for which j^n consists at the first k-1 positions the sequence $(i_1, i_2, ..., i_{k-1})$ (let say x_k^m),
3. Assign the recognized pattern into the class j_k^m.

C. The Fuzzy Methods

Now we take to decision algorithms for the MC using the inference engine that makes inferences on a fuzzy rule system. For all the algorithms presented below there is a common rule form for rules that associate an observation vector $a = (a^{(1)}, a^{(2)}, ..., a^{(n)})$ with a diagnosis. Further, we assume the following general form of the *k*-th rule in the system ($k = 1, 2, ..., K$):

$$IF\ a^{(1)}\ is\ A_{1,k}\ AND ... AND\ a^{(n)}\ is\ A_{n,k}\ THEN\ b\ is\ B_k, \quad (5)$$

where $A_{i,k}$ are fuzzy sets (whose membership functions are designated by $m_{A_{i,k}}$) that correspond to the nature of particular observations (for simplicity we assume the sets to be triangular fuzzy numbers) whereas B_k is a discrete fuzzy set defined on the diagnosis set, with the m_{B_k} membership function.

The particular decision algorithms to be used in MC have in common both the inference engine and the procedure for rule system (5) derivation from the learning set (4). The only difference between the algorithms is the form of observation vector *a* and diagnosis *b*, what follows, the number and nature of premises and conclusion in the rules (5). Let us first briefly present both of the common procedures.

The Inference Engine Procedure—composed of the three following steps:

1. For the *k*-th rule we determine the *degree of fulfilment* factor for a given observation *a*:

$$n_k(a) = \min_{i=1,2,...,L} m_{i,k}(a_i). \quad (6)$$

2. Combination of fuzzy rules — we determine the resultant consequence in the form of a fuzzy set *B* using the principle of maximum combination of responses:

$$m_B(x, a) = \max_{k=1,2,...,L} n_k(a) m_{B_k}(x). \quad (7)$$

3. Defuzzification — as we deal with discrete conclusions, we use defuzzification by maximum, namely:

$$j(a) = \max_x m_B(x, a). \quad (8)$$

A Procedure for Deriving the Rule Systems from Training Sets. We have accepted here the modified *b-cut algorithm* [4]that is composed of the following steps:

1. Find extremes for each variable in the training set ($l = 1,2,...,n$):

$$a^{(l)}_{\max} = \max_{i=1,2,...,L} a^{(l)}_i , a^{(l)}_{\min} = \min_{i=1,2,...,L} a^{(l)}_i$$

2. Find the triangular fuzzy number $(a^{(l)-}_k, a^{(l)}_k, a^{(l)+}_k)$ for the *l*-th variable and the *k*-th rule corresponding to the diagnosis $k \in M$:

$$a^{(l)-}_k = \min_{i=1,2,...,L} \{a^{(l)}_i \text{ such that } j^i = k\},$$

$$a^{(l)+}_k = \max_{i=1,2,...,L} \{a^{(l)}_i \text{ such that } j^i = k\},$$

$$a^{(l)}_k = \frac{1}{N_k} \sum \{a^{(l)}_i \text{ such that } j^i = k\},$$

where N_k is a number of such elements that $j^i = k$.

3. Calculate for each variable the relative diameter of the support:

$$d_l = \min_{k \in M} \frac{a^{(l)+}_k - a^{(l)-}_k}{a^{(l)}_{\max} - a^{(l)}_{\min}}.$$

Variables with $d_l > \Delta$ are not considered to belong to the rule system. The other ones will constitute the selected set of *n* explanatory premises.

<u>The fuzzy algorithm of MC for direct approach (Fuzzy A)</u>

Now we have one system of fuzzy rules, where $a = x$ and $b = (j_1, j_2,..., j_N)$.

<u>The fuzzy algorithm of MC for decomposed independent approach (Fuzzy I)</u>

In this approach there are *N* sets of fuzzy rules for each task of MC. In rules for *k*-th task $a = x_k$ and $b = j_k$.

<u>The fuzzy algorithm of MC for decomposed dependent approach (Fuzzy D)</u>

As previously, but now $a = (x_k, j_1,..., j_{k-1})$ and $b = j_k$.

D. The Neural Network Approach

Similarly to the fuzzy approach, applying artificial neural networks as an implementation of the decision algorithm for MC is concerned exclusively with the relevant selection of input data.

Two kinds of neural networks have been accepted for the needs of a comparative analysis, namely the Back Propagation (BP) and Counter Propagation (CP) networks [5].

The Type BP Network

The network consists of the input, hidden, and output neuron layers. The input layer plays the role of a data buffer so that the data are normalized to belong to the [0, 1] range. There have been various numbers of input layer neurons, depending on actual quantities of data. Both the hidden and output layer neurons have the sigmoid transition function; the corresponding layers are trained by means of the error back propagation method. Neurons of the successive layers are connected on the each-to-each basis.

The Type CP Network

The network has three layers of neurons. Just as it was the case in the BP network, the input layer normalizes input vectors to remove analysis errors that are typical for the CP networks. The second layer, trained according to the Kohonen algorithm, possesses an empirically estatblished number of neurons that depends on input data quantities and ranges within the limits from 50 to 70 neurons. The output layer has been trained using Grossberg's *outstar* method. Neuron layers are connected on the each-to-each basis.

The input data sets are just the same as those for the fuzzy-approach algorithms. Thus the BP-A and CP-A designations correspond to type BP and CP networks, respectively, for direct approach of MC, ie. with input vector *x* and output $(j_1, j_2,..., j_N)$. Further, BP-I, CP-I, BP-D and CP-D algorithm designations denote the relevant networks used for decomposed independent and decoposed dependent approaches, respectively.

All the decision algorithms that are depicted in this chapter have been experimentally tested as far as the decision quality is concerned. Measure for the decision quality is the frequency of correct diagnoses for real data that are concerned with multiperspective recognition of chronic renal failure and decisions in non-Hodgkin lymphoma.. The purpose of our research and associated tests was not only the comparative analysis of the presented algorithms but also answering the question whether decomposed dependent approach would yield a better decision quality as compared to decomposed independent method. The next chapter describes the performed tests and their outcome.

IV. PRACTICAL EXAMPLES

A. Diagnosis of chronic renal failure in children

Chronic renal failure (CRF) is a syndrome of of clinical symptoms caused by the adverse action of a factor or factors of the urinary tracts. It is characterized by the failure to eliminate the final products of nitrogen metabolism, acid-base imbalance, failure to maintain the water-electrolyte balance and a damage to the endocrine function of the kidneys.

The complete diagnosis of CRF contains the two following task:

- recognition of etiologic type of CRF (glomerulonephritis, pyelonephritis, renal calculosis, renal defect, metabolic disorders, CRF as a result of systemic disease),
- recognition of CRF phase (decadent, compensated, and decomensated CRF).

The vector of features contains the values of 53 clinical data presented in Table 1.

In the Department of Pediatric Nephrology the set of 380 case records of children suffering from chronic renal failure were collected, which constitute the learning set (4). Results

of experimental investigations of MC algorithms for different approaches are presented in Table 3.

TABLE 1
CLINICAL FEATURES CONSIDERED

GENERAL: Age, Sex
PAST AND/OR CONCURRENT DISEASES: Organic heart disease, Nephrosis, Surgery (abdominal), Abdominal tumor, Streptococcal pharyngitis, Scarlatina, Urinary tracts lithiasis
ANAMNESIS: Fever, Hereditary transmission, Diarrhoea, Vomiting, Convulsions, Growth of signs
PHYSICAL EXAMINATIONS: Condition of child, Skin, Blood pressure (syst. and diast.), Extravasation, Oedema, Puls e, Body temperature, Urine in bladder
LABORATORY EXAMINATIONS: Sedimentation rate
GASOMETRIC EXAMINATIONS: p O_2, p CO_2, pH, BE
MORPHOLOGY: Leucocytes, Reticulocytes, Trombocytes, Erytrocytes, Hemoglobin
SERUM: Urine level, Total protein level
SERUM IONOGRAM: Na^+, K^+
URINE: Micturition, Intermittent diuresis, Haematuria, Erythrocyturia, Erythrocytes, Protein, Cylinders
OTHERS: Pain, Urine culture, Toxins, Arrhythmia, Inflammation of heart muscle, Infection of respiratory tracts, Dyspnoea, Pulmonary oedema, Dehydration

B. Decision problems in non-Hodgkin lymphoma (NHL)

The non-Hodgkin lymphoma is a common dilemma in hematologic practice. It provides a physician with many difficult decision problems which can be treated as a multiperspective recognition presented in Table 2.

TABLE 2
MULTIPERSPECTIVE RECOGNITION IN NON-HODGKIN LYMPHOMA

Task	Decision	Features
1	Lymphoma type (small degree of malignancy, moderate degree of malignancy, immunoblastic lymphoma, lymphoblastic lymphoma, Burkitt-type lymphoma)	cytologic examinatons of cells, lymphoma structure (e.g. diffuse, nodulous, etc.)
2	Stage of lymphoma development (first, second, third and fourth stage of lymphoma development)	Nodes, tissues and organs affected
3	Lymphoma form (form A – no additional features, form B – with additional features)	Additonal features (fever, insomnia, diaphoresis, etc.)
4	Scheme of treatment CHOP, BCVP, COMBA, MEVA, COP, CHOP-BLEO	Decisions i_1, i_2, and i_3
5	Modification of scheme (recommendations for grading of acute and subacute toxicity)	Side effects (hair, infection, cardiac rhythm, pain, etc.)

In order to study the performance of the proposed recognition concepts and evaluate their usefulness to the computer-aided decisions in non-Hodgkin lymphoma some computer experiments were made using the set of 225 case records collected in the Wroclaw Medical Academy. The outcome is shown in Table 3.

V. CONCLUSION

Results of experiments imply the following conclusions:
1. Out of all the above-mentioned approaches to multiperspective medical diagnosis, the best outcome is that achieved as a result of using the type Back Propagation neural network. The probabilistic algorithm yields a little worse results. Still worse are the results obtained from the Counter Propagation neural network whilest the fuzzy logic algorithms turn out to be the undoubtedly worst ones.
2. There occurs a common effect within each algorithm group: algorithms for decomposed independent approach are always worse than rules for direct and decomposed dependent concept.

The comparative analysis presented above for the multiperspective diagnosis algorithm is also of the experimental nature. We have carried out a series of experiments on the basis of a specific exemplar that concerns chronic renal failure and non-Hodgkin lymphoma diagnosing using rich enough set of real-life data. The objective of our experiements was to measure quality of the tested algorithms that was defined by the frequency of correct decisions. The algorithm-ranking outcome cannot be treated as one having the ultimate character as a that of a law in force, but it has been achieved for specific data within a specific diagnostic task. However, although the outcome may be different for other tasks, the presented research may nevertheless suggest some perspectives for practical applications.

TABLE 3
RESULTS OF ACCURACY OF COMPUTER-AIDED DIAGNOSIS OF CHRONIC RENAL FAILURE (CRF) AND DECISIONS IN NON-HODGKIN LYMPHOMA (NHL) FOR DIFFERENT ALGORITHMS AND CONCEPTS OF MULTIPERSPECTIVE RECOGNITION

Approach	Algorithm	CRF Accuracy [%]	NHL Accuracy [%]
Direct approach	Bayes - A	90,3	83,9
	NN-A	80,8	78,3
	Fuzzy - A	71,3	74,2
	BP - A	92,3	85,3
	CP-A	85,7	81,2
Decomposed independent approach	Bayes - I	71,4	68,4
	NN - I	65,3	59,3
	Fuzzy - I	55,3	50,8
	BP - I	73,8	69,3
	CP - I	68,4	66,5
Decomposed dependent approach	Bayes - D	86,4	81,6
	NN - D	75,6	76,3
	Fuzzy - D	69,4	70,5
	BP - D	89,6	84,1
	CP - D	82,6	80,9

REFERENCES

[1] M.Ben-Bassat, Multimembership and multiperspective classification, *IEEE Trans on SMC*, vol.10, pp.331-336, 1980.
[2] M. Kurzynski, E. Puchala, Algorithms of the Multiperspective Pattern Recognition, *Proc 11th Int. Conference on Pattern Recognition*, vol. B, pp. 627-631, Hague 1992
[3] L. Devroye, L. Gyorfi, *A Probabilistic Theory of Pattern Recognition*, Springer Verlag, New York-Berlin-Tokyo 1996
[4] J. Bezdek, S. Pal [eds.], *Fuzzy Models for Pattern Recognition*, IEEE Press, New York 1992
[5] B. D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge University Press, 1996